

日本国特許庁
JAPAN PATENT OFFICE

#5
CERTIFIED COPY OF
PRIORITY DOCUMENT
3404

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日
Date of Application:

2001年10月30日

出願番号
Application Number:

特願2001-331858

[ST.10/C]:

[JP2001-331858]

出願人
Applicant(s):

株式会社日立製作所

USSN 10/076,456

MATTINGLY, STANGER + MALUR

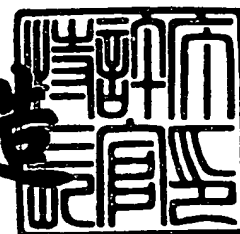
(703) 684-1120

DKT: NIT-329

2002年 3月 1日

長官
Office

及川耕造



出証番号 出証特2002-3011687

【書類名】 特許願

【整理番号】 K01009841A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

 【氏名】 吉田 晃

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

 【氏名】 中村 崇仁

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

【手数料の表示】

 【予納台帳番号】 013088

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 ディスクアレイ制御装置

【特許請求の範囲】

【請求項 1】

ホストコンピュータに接続されるホストスイッチインターフェース部と、

ホストスイッチインターフェース部とのインターフェースを有する複数のチャネルインターフェース部、磁気ディスク装置に接続される複数のディスクインターフェース部、及び、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納する複数のキャッシュメモリ部とを有する複数のディスクアレイ制御ユニットと、

前記複数のチャネルインターフェース部、前記複数のディスクインターフェース部、及び、前記複数のキャッシュメモリ部とに接続される相互結合網とを有するディスクアレイ制御装置であって、

前記キャッシュメモリは、すべての前記ディスクアレイ制御ユニットの前記チャネルインターフェース部とデータ転送を行なえることを特徴とするディスクアレイ制御装置。

【請求項 2】

前記ホストスイッチインターフェース部は、ホストコンピュータからのリクエストのアドレスに基づいてデータ転送経路を選択する管理テーブルを有することを特徴とする請求項 1 記載のディスクアレイ制御装置。

【請求項 3】

前記データ転送経路は、前記ホストスイッチインターフェース部と前記チャネルインターフェース部とのパスであることを特徴とする請求項 2 記載のディスクアレイ制御装置。

【請求項 4】

前記管理メモリは、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、データ転送経路ごとのデータ量に基づいて重み付けされた履歴情報テーブルを有し、前記パス選択テーブルで選択されたパスの中から前記履歴情報テーブルの情報に基づいてパスを選択することを特徴とする請求項 2 記載のディ

スクアレイ制御装置。

【請求項5】

前記複数のディスクアレイ制御ユニットのうち一部のディスクアレイ制御ユニットは、他のディスクアレイ制御ユニットより高速のキャッシュメモリを有することを特徴とする請求項1記載のディスクアレイ制御装置。

【請求項6】

前記複数のディスクアレイ制御ユニットは、それぞれ該ディスクアレイ制御ユニットのリソースの稼働率の管理を行ない、稼働率を稼働率報告信号によって前記ホストスイッチインターフェース部に報告するリソース管理部を有し、

前記管理メモリは、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、前記稼働率報告信号に基づいて重み付けされた履歴情報テーブルを有し、前記パス選択テーブルで選択されたパスの中から前記履歴情報テーブルの情報に基づいてパスを選択することを特徴とする請求項2記載のディスクアレイ制御装置。

【請求項7】

ホストコンピュータに接続されるホストスイッチインターフェース部と、

ホストスイッチインターフェース部とのインターフェースを有する複数のチャネルインターフェース部、磁気ディスク装置に接続される複数のディスクインターフェース部、及び、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納する複数のキャッシュメモリ部とを有する複数のディスクアレイ制御ユニットと、

前記複数のチャネルインターフェース部、前記複数のディスクインターフェース部、及び、前記複数のキャッシュメモリ部とに接続される第1の相互結合網と

前記ホストスイッチインターフェース部と前記複数のチャネルインターフェース部とに接続される第2の相互接続網とを有するディスクアレイ制御装置であって、

前記キャッシュメモリは、すべての前記ディスクアレイ制御ユニットの前記チャネルインターフェース部とデータ転送を行なえることを特徴とするディスクア

レイ制御装置。

【請求項 8】

前記ホストスイッチインターフェース部は、ホストコンピュータからのリクエストのアドレスに基づいてデータ転送経路を選択する管理テーブルを有することを特徴とする請求項 7 記載のディスクアレイ制御装置。

【請求項 9】

前記データ転送経路は、前記第 2 の相互接続網とチャネルインターフェース部とのパスであることを特徴とする請求項 8 記載のディスクアレイ制御装置。

【請求項 10】

前記管理メモリは、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、データ転送経路ごとのデータ量に基づいて重み付けされた履歴情報テーブルを有し、前記パス選択テーブルで選択されたパスの中から前記履歴情報テーブルの情報に基づいてパスを選択することを特徴とする請求項 8 記載のディスクアレイ制御装置。

【請求項 11】

前記複数のディスクアレイ制御ユニットのうち一部のディスクアレイ制御ユニットは、他のディスクアレイ制御ユニットより高速のキャッシュメモリを有することを特徴とする請求項 7 記載のディスクアレイ制御装置。

【請求項 12】

前記複数のディスクアレイ制御ユニットは、それぞれ該ディスクアレイ制御ユニットのリソースの稼働率の管理を行ない、稼働率を稼働率報告信号によって前記ホストスイッチインターフェース部に報告するリソース管理部を有し、

前記管理メモリは、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、前記稼働率報告信号に基づいて重み付けされた履歴情報テーブルを有し、前記パス選択テーブルで選択されたパスの中から前記履歴情報テーブルの情報に基づいてパスを選択することを特徴とする請求項 8 記載のディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ディスクアレイ制御装置に関し、特に、データを複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置技術に関する。

【0002】

【従来の技術】

今般のコンピュータシステムにおいて、処理性能の向上に対する期待は大きく、特にディスクサブシステムのI/O性能向上に対する要求は高い。磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシステム」という。）のI/O性能は半導体記憶装置を記憶媒体とするコンピュータの主記憶のI/O性能に比べて、3～4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上させる努力がなされている。

【0003】

また、銀行、証券、電話会社等に代表される大企業では、従来は各所に分散していたコンピュータおよびストレージを、データセンターの中に集中化してコンピュータシステムおよびストレージシステム構成することにより、コンピュータシステムおよびストレージシステムの運用、保守、管理に要する費用を削減する傾向にあり、特に大型／ハイエンドのストレージシステムには、数百台以上のホストコンピュータへ接続するためのチャネルインターフェースのサポート（コネクティビティ）、数百テラバイト以上の記憶容量のサポートが要求されている。

【0004】

一方、近年のオープン市場の拡大、今後予想されるストレージ・エリア・ネットワーク（SAN）の普及により、大型／ハイエンドのストレージシステムと同様の高機能・高信頼性を備えた小規模構成（小型筐体）のストレージシステムへの要求が高まっている。

【0005】

サブシステムのI/O性能を向上させるための1つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。ディスクアレイの場合、上位コンピュータからのI/Oを記録する複数の磁気ディスク装置

と、上位コンピュータのI/Oを受付け複数の磁気ディスク装置へ転送するディスクアレイ制御装置から構成されるのが一般的である。この中で、大規模接続・大容量への要求に対しては、従来の大型／ハイエンドのディスクアレイ制御装置を複数接続して超大規模なディスクアレイ制御装置を構成する方法が考えられる。複数のディスクアレイ制御装置を接続することにより、キャッシュメモリが複数のディスクアレイ制御装置に分散することになる。キャッシュメモリは当該記憶制御装置に繋がる磁気ディスク装置のデータを格納することが性能の面を考えると有利であり、また、該キャッシュメモリにアクセスするホストコンピュータは、当該キャッシュメモリを有する当該記憶制御装置内に繋がるのが性能の面で有利である。また、実装障害時および磁気ディスク装置の増設や記憶制御装置の増設により、ホストインタフェース部とキャッシュメモリおよびキャッシュメモリと磁気ディスク制御装置との対応が上記性能面で有利と考える構成から変更となる場合がある。上位装置およびソフトから見た場合に複数のディスクアレイ制御装置を複数接続した超大規模なディスクアレイ制御装置と意識すること無く、一つのディスクアレイ制御装置として管理することが従来のアーキテクチャ継続の面から有利である。このような事より、ホストインタフェースとキャッシュメモリや磁気ディスク装置の接続関係や増設時の再構成に伴い実装位置の最適化を考える必要があり、従来の一つのディスクアレイ制御装置によりシステムが構築されていた場合に比べ、ディスクアレイシステムの性能に与える影響が大きくなる。

【0006】

例えば、従来技術では、図2に示すようにホストコンピュータ50とディスクアレイ制御装置2との間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置5とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置5のデータを一時的に格納するキャッシュメモリ部14とを備え、1つのディスクアレイ制御装置2内において、キャッシュメモリ部14は全てのチャンネルIF部11及びディスクIF部12からアクセス可能な構成となっている。このディスクアレイ制御装置2では、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14との間は、相互

結合網 21 で接続される。

【0007】

チャンネル I F 部 11 は、ホストコンピュータ 50 と接続するためのインターフェース及びホストコンピュータ 50 に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスク I F 部 12 は、磁気ディスク装置 5 と接続するためのインターフェース及び磁気ディスク装置 5 に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスク I F 部 12 は、RAID 機能の実行も行う。

【0008】

この従来のディスクアレイ制御装置 2 では、1 つのディスクアレイ制御装置 2 がサポートする記録データ量以上のデータを記録する必要がある場合、ディスクアレイ制御装置 2 を複数台設置し、ホストコンピュータ 50 から複数のディスクアレイ制御装置 2 にチャンネルを接続していた。

【0009】

また、1 つのディスクアレイ制御装置 2 に接続できるホストチャンネル数以上のホストコンピュータ 50 を接続する必要がある場合、ディスクアレイ制御装置 2 を複数台設置し、それぞれにホストコンピュータ 50 を接続していた。

【0010】

また、複数のディスクアレイ制御装置 2 間でデータを移行させる場合、1 つのホストコンピュータ 50 からデータの移行を行う 2 つのディスクアレイ制御装置 2 の両方にチャンネルを接続し、ホストコンピュータ 50 を介して、データの移行を行っていた。

【0011】

また、他の従来技術では、図 3 に示すようにホストコンピュータ 50 とディスクアレイ制御装置 2 と、ホストコンピュータ 50 とディスクアレイ制御装置 2 とを跨る外部接続網 23 とホストコンピュータ 50 との間のデータ転送を実行する複数のチャンネル I F 部 11 と、磁気ディスク装置 5 とディスクアレイ制御装置 2 間のデータ転送を実行する複数のディスク I F 部 12 と、磁気ディスク装置 5 のデータを一時的に格納するキャッシュメモリ部 14 とを備え、ホストコンピュー

タ50は、外部接続網23を介し全てのディスクアレイ制御装置2にアクセスすることが可能であり、1つのディスクアレイ制御装置2内において、キャッシュメモリ部14は全てのチャンネルIF部11およびディスクIF部12からアクセス可能な構成となっている。このディスクアレイ制御装置2では、チャンネルIF部11およびディスクIF部12とキャッシュメモリ部14との間は、相互結合網21で接続される。

【0012】

チャンネルIF部11は、ホストコンピュータ50と接続するためのインターフェースおよびホストコンピュータ50に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、磁気ディスク装置5と接続するためのインターフェースおよび磁気ディスク装置5に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、RAID機能の実行も行う。

【0013】

【発明が解決しようとする課題】

図2に示す従来技術では、ディスクアレイ制御装置2の台数を単純に増やすことによってホストコンピュータ50に接続するチャンネル数、記憶容量を増やすことが可能であるが、1つのホストコンピュータ50が複数のディスクアレイ制御装置2にデータを格納する必要がある場合、ホストコンピュータ50は全てのディスクアレイ制御装置2に対してチャンネルを接続する必要があり、ホストコンピュータ50はアクセスするデータがどのディスクアレイ制御装置2に接続された磁気ディスク装置5に格納されているかを把握し、データにアクセスする場合、目的のディスクアレイ制御装置2を特定してアクセスを行う必要があった。

【0014】

図3に示す従来技術では、複数のディスクアレイ制御装置2をホストコンピュータ50から見て一つのディスクアレイ制御装置として運用する為には、外部接続網23内に外部接続網23に接続される全てのディスクアレイ制御装置2の情報を管理運用する手段を持つことが必要である。ディスクアレイ制御装置の使用法として、磁気ディスク装置内のデータを他の磁気ディスク装置にコピーするこ

とがよく行われる。図3に示す従来技術においては、あるディスクアレイ制御装置2に繋がる磁気ディスク装置5のデータを他のディスクアレイ制御装置2に繋がる磁気ディスク装置5にコピーをする場合、チャンネルインターフェース部11を介し、外部接続網23を介することにより、データ転送を行なうこととなる。

【0015】

【課題を解決するための手段】

上記課題を解決するため、本発明におけるディスクアレイ制御装置は、ホストコンピュータに接続されるホストスイッチインターフェース部と、ホストスイッチインターフェース部とのインターフェースを有する複数のチャンネルインターフェース部、磁気ディスク装置に接続される複数のディスクインターフェース部、及び、磁気ディスク装置に対しリード／ライトされるデータを一時的に格納する複数のキャッシュメモリ部と有する複数のディスクアレイ制御ユニットと、複数のチャンネルインターフェース部、複数のディスクインターフェース部、及び、前記複数のキャッシュメモリ部とに接続される相互結合網とを有する。そして、キャッシュメモリは、すべてのディスクアレイ制御ユニットのチャンネルインターフェース部とデータ転送を行なえる。また、前記ホストスイッチインターフェース部と複数のチャンネルインターフェース部とは、相互接続網を介して接続されるようにしてもよい。

【0016】

ホストスイッチインターフェース部は、ホストコンピュータからのリクエストのアドレスに基づいてデータ転送経路を選択する管理テーブルを有する。発明の実施の形態で説明するように、本発明の実施例では、データ転送経路はホストスイッチインターフェース部とチャンネルインターフェース部とのパスである。

【0017】

管理メモリは、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、データ転送経路ごとのデータ量に基づいて重み付けされた履歴情報テーブルを有し、パス選択テーブルで選択されたパスの中から履歴情報テーブルの情報に基づいてパスを選択する。

【0018】

また、ディスクアレイ制御ユニットに、それぞれディスクアレイ制御ユニットのリソースの稼働率の管理を行ない、稼働率を稼働率報告信号によって前記ホストスイッチインターフェース部に報告するリソース管理部を設け、管理メモリが、アドレスに対するデータ転送経路の候補を有するパス選択テーブルと、稼働率報告信号に基づいて重み付けされた履歴情報テーブルを有し、パス選択テーブルで選択されたパスの中から前記履歴情報テーブルの情報に基づいてパスを選択するようにしてもよい。

【0019】

その他、本願が開示する課題、およびその解決方法は、発明の実施形態の欄および図面により明らかにされる。

【0020】

【発明の実施の形態】

以下、本発明の実施例を図面を用いて説明する。

〔実施例1〕

図1および図4に、本発明の第1の実施例を示す。

【0021】

図1に示すように、ディスクアレイ制御装置1は複数のディスクアレイ制御ユニット1-2およびホストスイッチインターフェース部（ホストスイッチIF部）30から構成される。ディスクアレイ制御ユニット1-2は、ホストスイッチIF部30とのインターフェース部（チャンネルIF部）11と、磁気ディスク装置5とのインターフェース部（ディスクIF部）12と、キャッシュメモリ部14を有し、チャンネルIF部11およびディスクIF部12とキャッシュメモリ部14間には、ディスクアレイ制御ユニット1-2内において相互結合網21を介して接続されている。また、複数のディスクアレイ制御ユニット1-2の間では、キャッシュメモリ部14は相互結合網21を介して接続されている。すなわち、相互結合網21を介して、全てのチャンネルIF部11およびディスクIF部12から、全てのキャッシュメモリ部14へアクセス可能な構成となっている。相互接続網21は、ディスクアレイ制御ユニット内のデータ転送性能の方が複数のディスクアレイ制御ユニットを介するデータ転送性能より高くなっている。即ち、ホ

ストコンピュータからのアクセスでは、アクセス対象のキャッシュメモリと同じディスクアレイ制御ユニット内のチャンネルＩＦ部を介してデータ転送を行なったほうが相互接続網への負担は小さくなる。また、ホストスイッチＩＦ部３０と複数のディスクアレイ制御ユニット１－２内の複数のチャンネルＩＦ部１１の間は、ホストスイッチＩＦ部３０から、全てのキャッシュメモリ部１４へアクセス可能な構成となっている。ホストスイッチＩＦ部３０は、ホストコンピュータからの要求に対して、すべての複数のディスクアレイ制御ユニット１－２内の複数のチャンネルＩＦ部１１からどのチャンネルＩＦ部に接続するか選択することができる。ホストスイッチＩＦ部３０内には、ホストアクセス情報および接続情報を格納する管理テーブル３１を持つ。

【００２２】

図１にて、ホストコンピュータ５０からのディスクアレイ制御ユニット１－２内のキャッシュメモリ部１４への読み出し処理について考える。ホストスイッチＩＦ部３０は、ホストコンピュータ５０からのリクエスト（リード／ライト要求）に対し、ホストスイッチＩＦ部３０内にある管理テーブル３１を参照し、対象キャッシュメモリ部１４にアクセス可能な最適ルートを求め、最適ルートが接続されるチャンネルＩＦ部１１に対しリクエストを送出する。リクエストを受け取ったチャンネルＩＦ部１１は、アドレス情報より相互結合網２１を介し、対象キャッシュメモリ部１４にアクセスを行なうことにより要求データを読み出す。ホストスイッチＩＦ部３０は、ホストコンピュータ５０からのリクエストをチャンネルＩＦ部１１に送出した時点で、アクセスのデータ量を元に重みづけを行ない管理テーブル３１内の履歴情報テーブルに情報を格納する。

【００２３】

図４に示すように、ディスクアレイ制御装置１は複数のディスクアレイ制御ユニットとホストスイッチインターフェース部（ホストスイッチＩＦ部）３０を有している。図４では、複数のディスクアレイ制御ユニットとしてディスクアレイ制御ユニット（ＢＯＸ０）１－２－１およびディスクアレイ制御ユニット（ＢＯＸ１）１－２－２およびディスクアレイ制御ユニット（ＢＫＢＯＸ）１－２－３が示されている。図４では、ＢＫＢＯＸはバックアップ用のディスクアレイ制御

ユニットとして使用する構成としている。図4に示したディスクアレイ制御ユニットBOX0 1-2-1およびBOX1 1-2-2およびBKBOX 1-2-3は、ホストスイッチIF部30とのインターフェース部（チャンネルIF部）11と、磁気ディスク装置とのインターフェース部（図示せず）と、キャッシュメモリ部14を有し、チャンネルIF部11とキャッシュメモリ部14間は、複数のディスクアレイ制御ユニットの間で、相互結合網21を介して接続されている。ホストスイッチIF部30と複数のディスクアレイ制御ユニット内の複数のチャンネルIF部11の間は、パス0 50およびパス1 51およびパスBK 52を介して接続されている。ホストスイッチIF部30内には、管理テーブル31を有し、管理テーブル31内にはパス選択テーブル32および履歴情報テーブル33を有する。履歴情報テーブル33からパス選択テーブル32に対し、パス選択信号40が接続され、パス選択信号40によりパス番号41が選択される。

【0024】

図4により、ホストコンピュータ50からディスクアレイ制御ユニット（BOX0）1-2-1内のキャッシュメモリ部14に対するアクセスについてホストスイッチIF部30の処理を考える。ホストコンピュータ50からのリクエストがホストスイッチIF部30に到着することにより、管理テーブル31内のパス選択テーブル32を参照し、当該リクエストのアドレスから、リクエスト対象データを保持するキャッシュメモリ部14を実装するディスクアレイ制御ユニットが、ディスクアレイ制御ユニット（BOX0）1-2-1或いはディスクアレイ制御ユニット（BOX1）1-2-2或いはディスクアレイ制御ユニット（BKBOX）1-2-3のどれであるか知ることができる。ここでは、BOX0 1-2-1内のキャッシュメモリ部14にアクセスするリクエストの為、BOX0 1-2-1が選択される。パス選択テーブル32には、各キャッシュメモリ毎にホストスイッチIF部30からの各キャッシュメモリ部14へのアクセスルート情報をパス番号として登録しておくことにより、当該リクエストのアドレスからパス選択テーブル32を参照することにより、選択パス候補を知ることができる。ここでは、BOX0 1-2-1内のキャッシュメモリ部14にアクセス

するリクエストの為、候補1としてパス0 50、候補2としてパス1 51、候補3としてパスBK 52が登録されている。パス選択テーブル32に登録するパス候補の数は任意であり、本実施例では3パスとしている。パス0 50は、ホストスイッチIF部30とBOX0 1-2-1を接続するパスであり、パス1 51は、ホストスイッチIF部30とBOX1 1-2-2を接続するパスであり、パスBK 52は、ホストスイッチIF部30とBK BOX 1-2-3を接続するパスである。続いて、履歴情報テーブル33からのパス選択信号40により、パス0 50、パス1 51およびパスBK 52のいずれのパスを選択するかが決定する。パス選択信号40は、履歴情報テーブル33内に過去のリクエストのパス毎のデータ量を重みづけしたテーブルを持ち、候補毎の重みの差、重み一定以上、障害情報等を判断することにより生成する。パス選択信号40によりパス番号41が決まり、当該キャッシュメモリ部へのアクセスルートが決定する。パス0 50が選択された場合、ホストスイッチIF部30からパス0 50を介しBOX0 1-2-1内のキャッシュメモリ部14にアクセスすることになる。パス1 51が選択された場合、ホストスイッチIF部30からパス1 51を介しBOX1 1-2-2内のチャンネルIF部11を経由し、複数のディスクアレイ制御ユニットを跨る相互結合網21を介し、BOX0 1-2-1内のキャッシュメモリ部14にアクセスすることになる。パスBK 52が選択された場合、ホストスイッチIF部30からパスBK 52を介しBK BOX 1-2-3内のチャンネルIF部11を経由し、複数のディスクアレイ制御ユニットを跨る相互結合網21を介し、BOX0 1-2-1内のキャッシュメモリ部14にアクセスすることになる。履歴情報テーブル33の重みが均一の場合、最短でアクセスすることが可能であるパス0 50が選択される。

【0025】

図5および図6に、ホストスイッチIF部30と複数のディスクアレイ制御ユニット1-2内の複数のチャンネルIF部11の間を相互結合網22を介して接続した実施例を示す。

【0026】

図5に示すように、ディスクアレイ制御装置1は複数のディスクアレイ制御ユ

ニット1-2およびホストスイッチIF部30から構成される。ホストスイッチIF部30と複数のディスクアレイ制御ユニット1-2内の複数のチャンネルIF部11の間は、相互結合網22を介して接続されている。すなわち、相互結合網22を介して、ホストスイッチIF部30から、全てのキャッシュメモリ部14へアクセス可能な構成となっている。キャッシュメモリ部14間を接続する相互結合網21とホストスイッチIF部30と複数のディスクアレイ制御ユニット1-2内の複数のチャンネルIF部11を接続する相互接続網22は、各々独立に動作可能である。その他の部分は図1に示した実施例と同じである。

【0027】

図5でホストコンピュータ50からのディスクアレイ制御ユニット1-2内のキャッシュメモリ部14への読み出し処理を考えると、ホストスイッチIF部30は、ホストコンピュータ50からの要求に対し、ホストスイッチIF部30内にある管理テーブル31を参照し、対象キャッシュメモリ部14にアクセス可能な最適ルートを求め、最適ルートが接続されるチャンネルIF部11に対し、相互接続網22を介しリクエストを送出する。ホストスイッチIF部30からのリクエストが相互接続網22を介して送出手されることを除けば、図1の実施例の動作と同じである。

【0028】

図6では、管理テーブル31内のパス選択テーブルおよび履歴情報テーブルの構成及び制御はホストスイッチIF部30とディスクアレイ制御ユニットが相互接続網22を介して接続されることを除けば、図4の実施例と同じである。パス0 50は、相互接続網22とBOX0 1-2-1を接続するパスであり、パス1 51は、相互接続網22とBOX1 1-2-2を接続するパスであり、パスBK 52は、相互接続網22とBK BOX 1-2-3を接続するパスである。パス選択テーブル32には、各キャッシュメモリ毎に相互接続網22からの各キャッシュメモリ部14へのアクセスルート情報がパス番号として登録されている。したがって、パス選択テーブル32を参照することにより、リクエストのアドレスから選択パス候補を知ることができる。ここでは、BOX0 1-2-1内のキャッシュメモリ部14にアクセスするリクエストの為、候補1とし

てパス 0 5 0、候補 2 としてパス 1 5 1、候補 3 としてパス BK 5 2 が登録されている。パス選択テーブル 3 2 に登録するパス候補の数は任意であり、本実施例では 3 パスとしている。続いて、履歴情報テーブル 3 3 からのパス選択信号 4 0 により、パス 0 5 0、パス 1 5 1 およびパス BK 5 2 のいずれのパスを選択するかが決定する。パス選択信号 4 0 によりパス番号 4 1 が決まり、当該キャッシュメモリ部へのアクセスルートが決定する。パス 0 5 0 が選択された場合、ホストスイッチ I F 部 3 0 から相互接続網 2 2 を介しパス 0 5 0 を使用して BOX 0 1 - 2 - 1 内のキャッシュメモリ部 1 4 にアクセスすることになる。パス 1 5 1 が選択された場合、ホストスイッチ I F 部 3 0 から相互接続網 2 2 を介しパス 1 5 1 を使用して BOX 1 1 - 2 - 2 内のチャンネル I F 部 1 1 を経由し、複数のディスクアレイ制御ユニットを跨る相互結合網 2 1 を介し、BOX 0 1 - 2 - 1 内のキャッシュメモリ部 1 4 にアクセスすることになる。パス BK 5 2 が選択された場合、ホストスイッチ I F 部 3 0 から相互接続網 2 2 を介しパス BK 5 2 を使用し BK BOX 1 - 2 - 3 内のチャンネル I F 部 1 1 を経由し、複数のディスクアレイ制御ユニットを跨る相互結合網 2 1 を介し、BOX 0 1 - 2 - 1 内のキャッシュメモリ部 1 4 にアクセスすることになる。

〔実施例 2〕

図 7 に、本発明の他の実施例を示す。

【0029】

図 7 の構成は、ディスクアレイ制御ユニット (BK BOX) 1 - 2 - 3 を通常動作時にも使用するディスクアレイ制御ユニットとして使用することとしており、BK BOX 1 - 2 - 3 内に高速キャッシュメモリ部 1 5 を有している。その他の構成については、実施例 1 に説明した図 4 の構成と同様である。図 5 により、高速なアクセスを要求されるデータ転送に対するホストスイッチ I F 部 3 0 の処理を考える。ホストコンピュータ 5 0 からのリクエストがホストスイッチ I F 部 3 0 に到着すると、管理テーブル 3 1 内のパス選択テーブル 3 2 を参照する。そのリクエストのアドレスから、リクエスト対象データは、高速キャッシュメモリ部 1 5 に実装されていることを知ることができる。履歴情報テーブル 3 3 か

らのパス選択信号40により、パスBK 52のパスが選択され、ホストスイッチIF部30からパスBK 52を介しBK BOX1-2-3内の高速キャッシュメモリ部15にアクセスすることになる。このように、特定のディスクアレイ制御ユニット内に高速なキャッシュメモリをもち、ホストスイッチIF部30内の管理テーブル31を有効に使用することにより、性能を要求されるホストからのリクエストに対する要求を満たすことが可能となる。もちろん、キャッシュメモリの容量を変えたり、パスを高速にすることも可能である。また、図4の実施例に対する図6の実施例のように、ホストスイッチIF部30とディスクアレイ制御ユニットとの間を相互接続網を介して接続する構成とすることもできる。この場合、上述したように、パス選択テーブル32に登録され履歴情報テーブル33に基づいて選択されるパスが相互接続網22とディスクアレイ制御ユニットとの間のものになる。

〔実施例3〕

図8に、本発明の他の実施例を示す。

【0030】

図8に示すように、複数のディスク制御ユニット内にリソース管理部16をもつこと、およびホストスイッチIF部30内の管理テーブル31内に稼働率管理テーブル34をもつこと、および複数のディスク制御ユニット内のリソース管理部16よりホストスイッチIF部30内の管理テーブル31内の稼働率管理テーブル34に対し稼働率報告信号53を接続していることを除いて、実施例1に説明した図4の構成と同様である。各々のディスク制御ユニット内のリソース管理部16は、各ディスク制御ユニット内のキャッシュメモリ部14およびチャンネルIF部11およびディスクIF部（図示せず）および複数のディスクアレイ制御ユニット間に跨る相互結合網21、および各ディスクユニット内の内部パスおよび内部バッファ等のリソースについての稼働率について管理を行なう。管理情報を稼働率報告信号53により、ホストスイッチIF部30内の管理テーブル31内の稼働率管理テーブル34に報告する。稼働率管理テーブル34は、各々のディスクアレイ制御ユニットからの稼働率報告信号53の内容について、各パス毎に重み付けを行ない管理する。その他の動作については、実施例1に記述した動

作と同様である。また、図4の実施例に対する図6の実施例のように、ホストスイッチIF部30とディスクアレイ制御ユニットとの間を相互接続網を介して接続する構成とすることもできる。この場合、上述したように、パス選択テーブル32に登録され履歴情報テーブル33に基づいて選択されるパスが相互接続網22とディスクアレイ制御ユニットとの間のものになる。

【0031】

本実施例によれば、稼働率報告信号53により詳細な稼働状況を知ることができ、より正確なパス選択を行なうことが可能となる。

【0032】

【発明の効果】

本発明によれば、複数台のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用しようとする場合、複数のディスクアレイ制御装置にキャッシュメモリが分散されて搭載された場合の、ホストコンピュータとキャッシュメモリ間での最適な接続ルートを確保することができ、リソースを最適に活用することが可能となる。また、ホストコンピュータにキャッシュメモリの物理的な実装位置を意識させること無く、1つのディスクアレイ制御装置としてキャッシュメモリの物理的な実装位置にかかわらず、一定したキャッシュメモリアクセスを提供することが可能となる。また、複数のディスクアレイ制御装置内のキャッシュメモリ部間を跨る相互結合網と、ホストインタフェースから複数のディスクアレイ制御装置を跨る相互結合網が独立に動作可能な為、複数のディスクアレイ制御装置間のキャッシュメモリ部間コピー動作により、キャッシュメモリ部間の相互結合網が専有された場合でも、ホストインタフェースからのリクエストは影響を受けること無く、各々のディスクアレイ制御装置に振り分けることが可能である。

【図面の簡単な説明】

【図1】

本発明によるディスクアレイ制御装置の構成を示す図。

【図2】

従来のディスクアレイ制御装置の構成を示す図。

【図 3】

従来のディスクアレイ制御装置の他の構成を示す図。

【図 4】

図 1 に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図 5】

本発明によるディスクアレイ制御装置の構成を示す図。

【図 6】

図 5 に示すディスクアレイ制御ユニット内の詳細構成を示す図。

【図 7】

本発明によるディスクアレイ制御装置の他の構成を示す図。

【図 8】

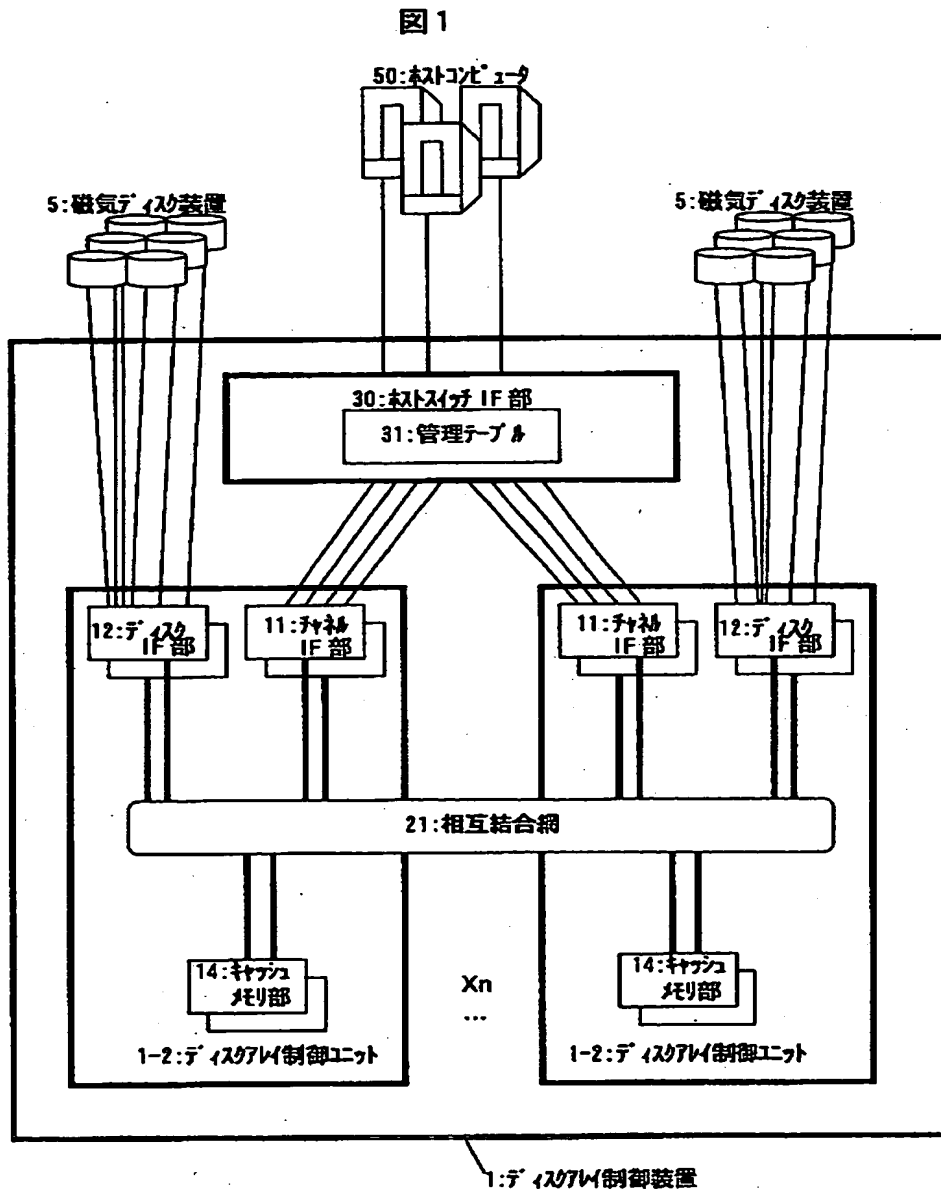
本発明によるディスクアレイ制御装置の他の構成を示す図。

【符号の説明】

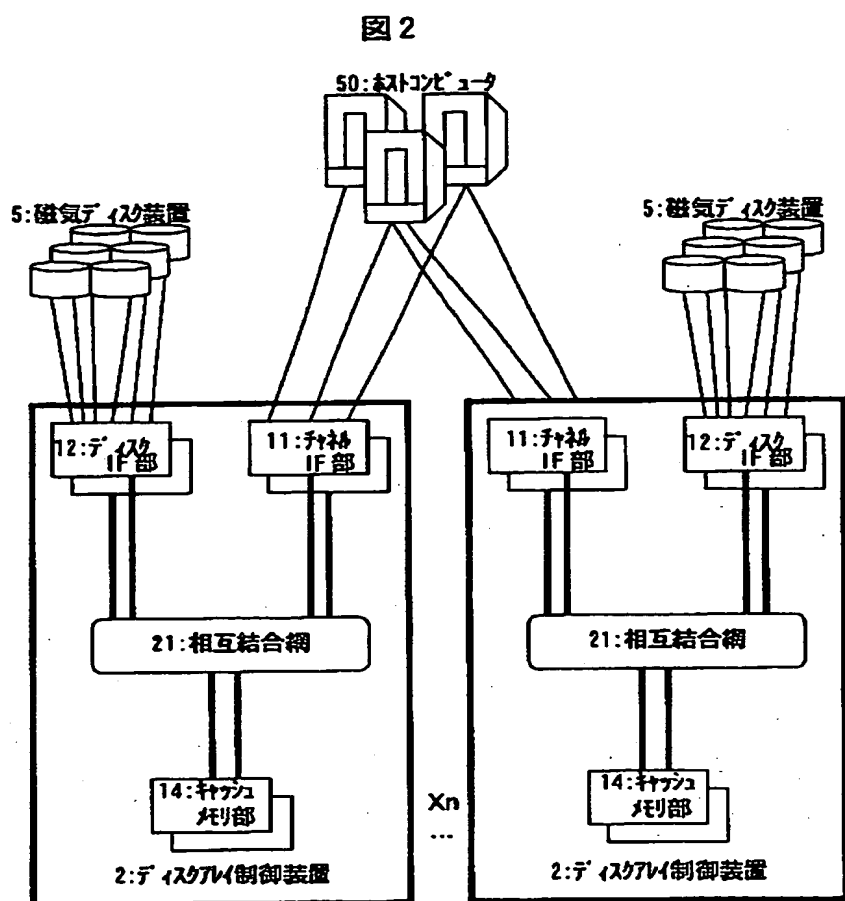
1, 2 : ディスクアレイ制御装置、 1-2 … ディスクアレイ制御ユニット、 5 … 磁気ディスク装置、 11 … チャンネル I/F 部、 12 … ディスク I/F 部、 14 … キャッシュメモリ部、 21, 22 … 相互結合網、 23 … 外部結合網、 30 … ホストスイッチ I/F 部、 31 … 管理テーブル、 50 … ホストコンピュータ

【書類名】 図面

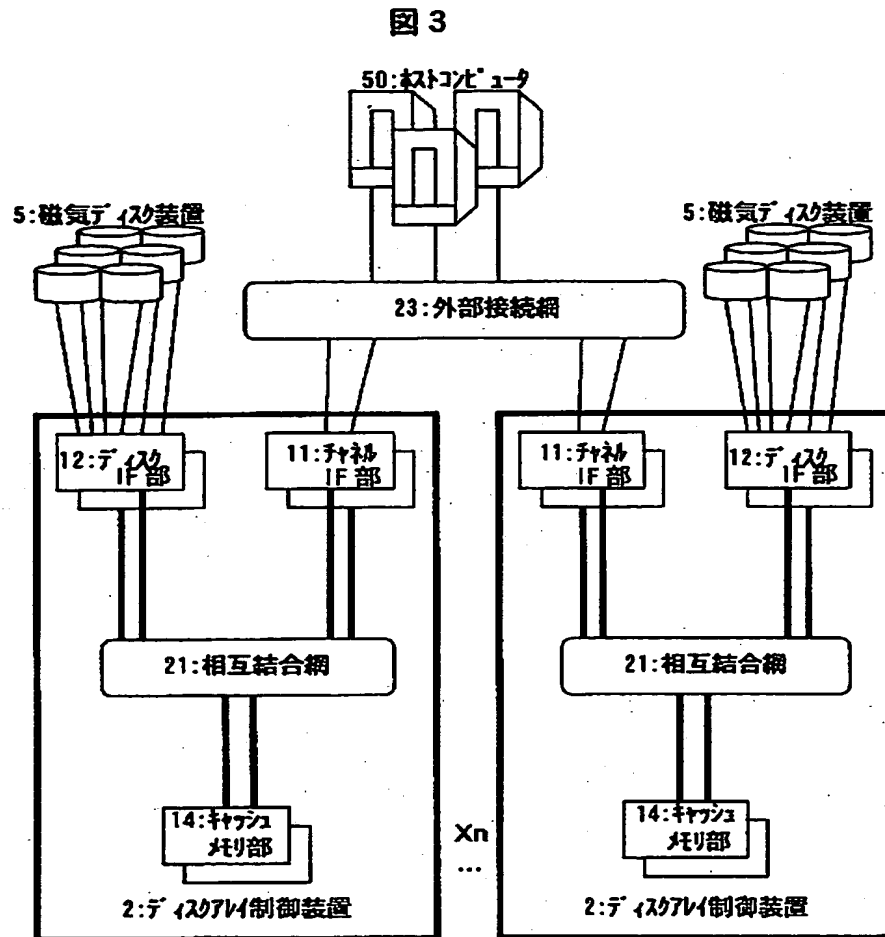
【図 1】



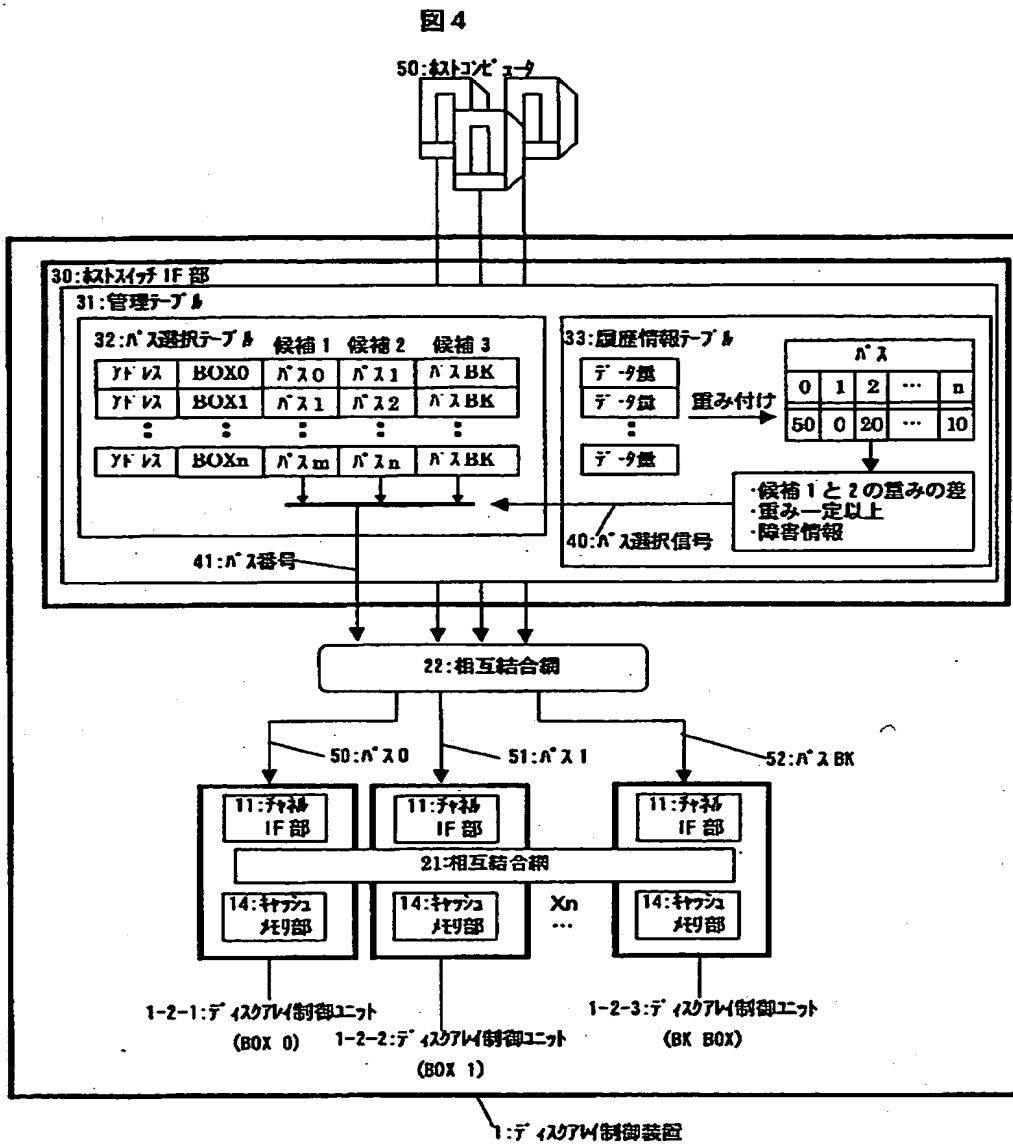
【図 2】



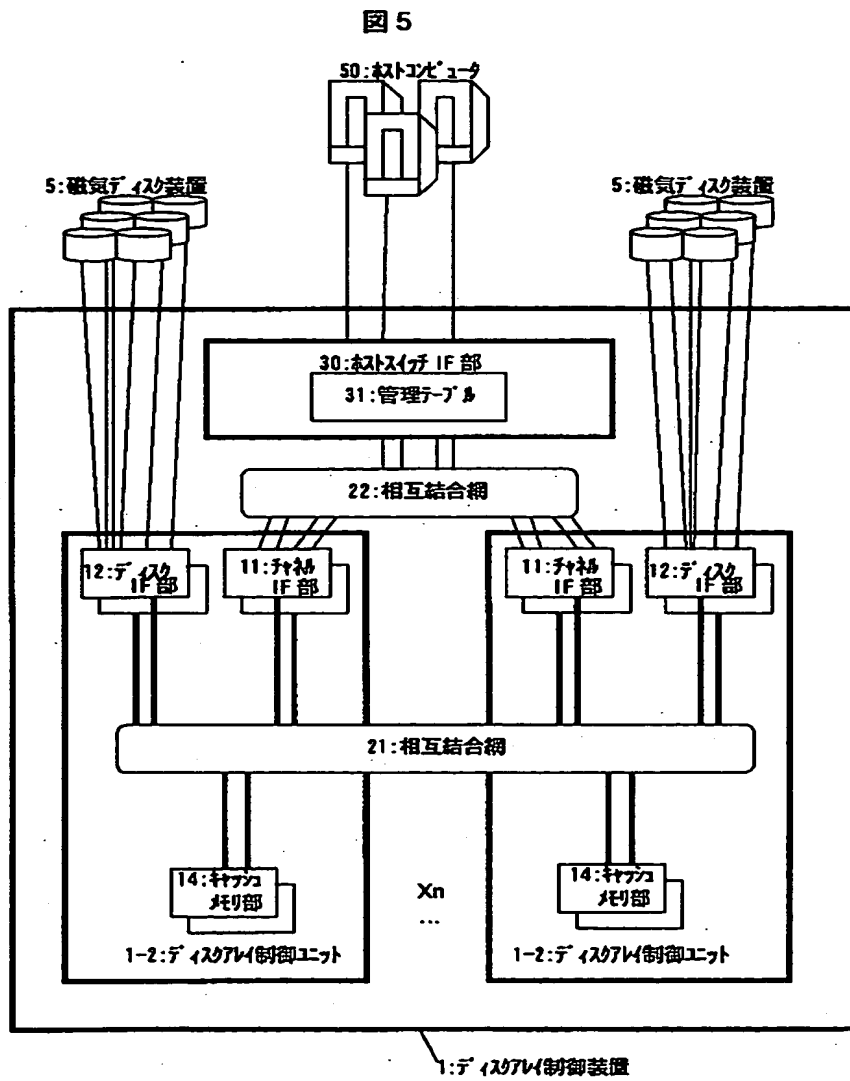
【図 3】



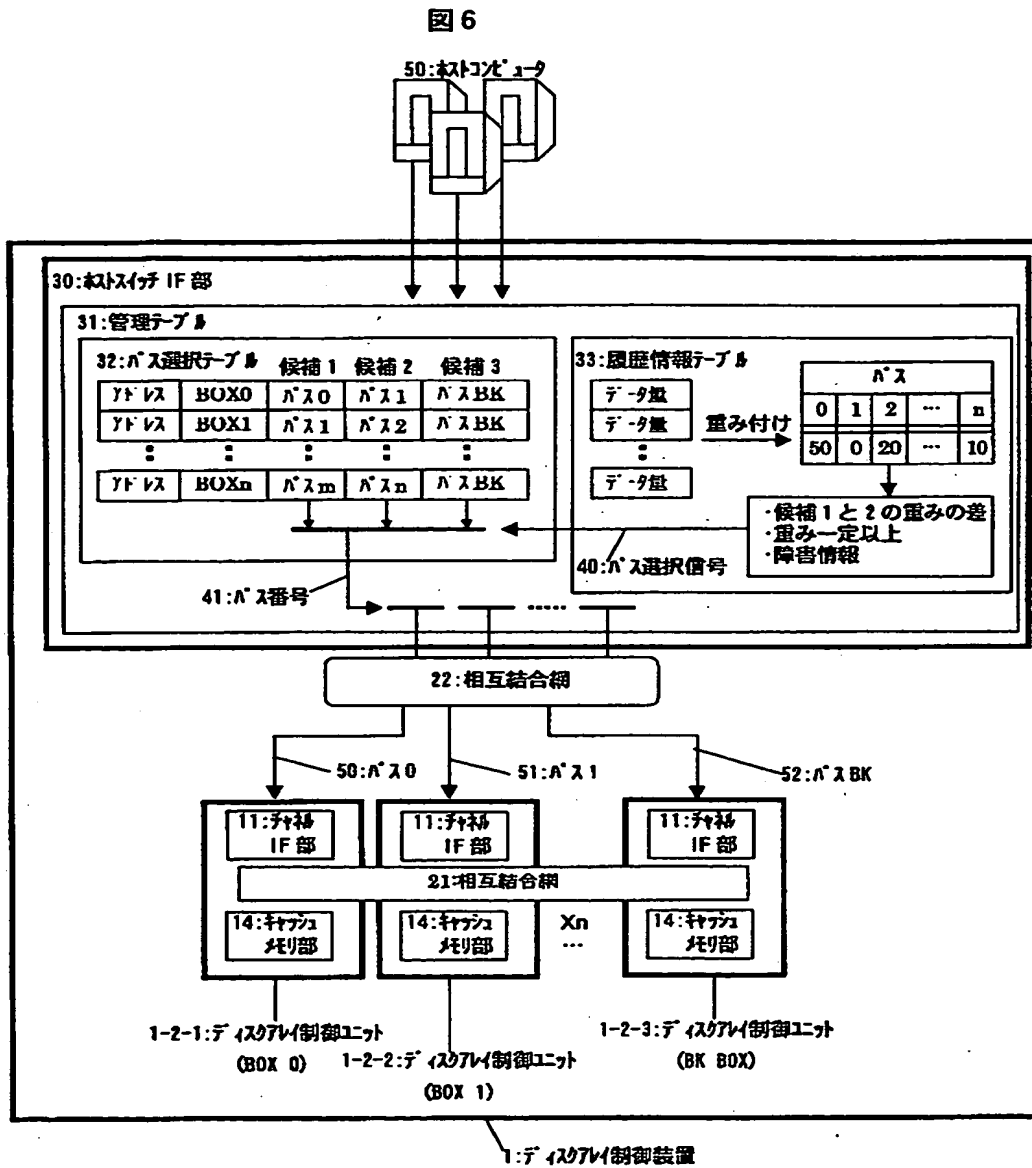
【図 4】



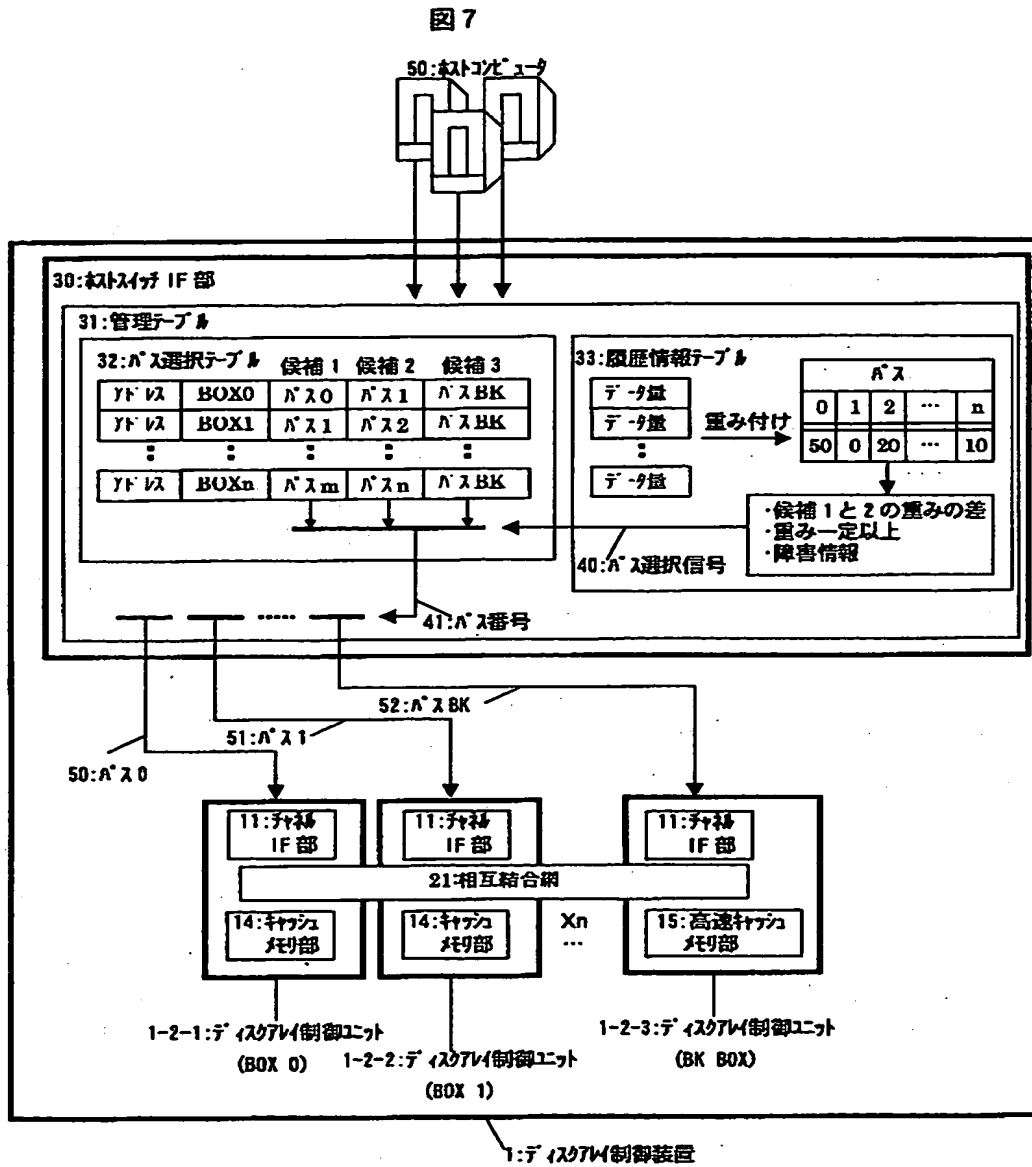
【図 5】



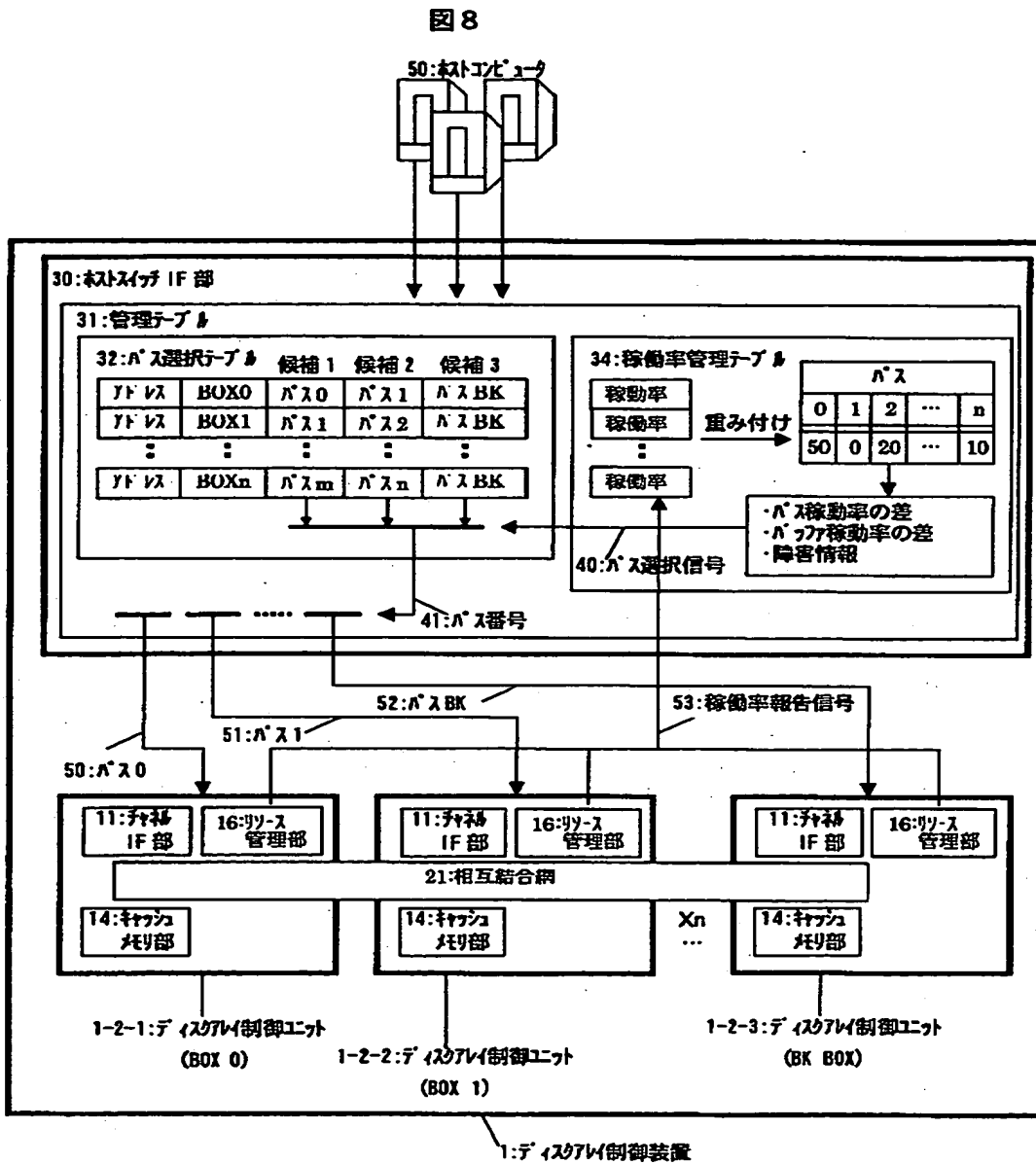
【図 6】



【図 7】



【図8】



【書類名】 要約書

【要約】

【課題】

複数のディスクアレイ制御装置を1つのディスクアレイ制御装置として運用でき、複数のディスクアレイ制御装置のキャッシュメモリの物理的な実装位置による性能低下を抑え、台数に比例した性能を出せるディスクアレイシステムを提供することにある。

【解決手段】

上記課題は、ホストスイッチインターフェース部と、複数のチャネルインターフェース部、複数のディスクインターフェース部、及び、複数のキャッシュメモリ部と有する複数のディスクアレイ制御ユニットと、複数のチャネルインターフェース部、複数のディスクインターフェース部、及び、前記複数のキャッシュメモリ部とに接続される相互結合網とを有するディスクアレイ制御装置により達成される。

【効果】

複数のディスクアレイ制御ユニット間に分散配置されているキャッシュメモリに対するアクセス性能が向上し、ディスクアレイ制御ユニットの台数に比例してディスクアレイ制御装置の性能を向上できる

【選択図】 図1

認定・付加情報

特許出願の番号	特願2001-331858
受付番号	50101597599
書類名	特許願
担当官	第七担当上席 0096
作成日	平成13年10月31日

<認定情報・付加情報>

【提出日】	平成13年10月30日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所